REPORT DOCUMENTATION PAGE	Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
	Technical Report	-

4. TITLE AND SUBTITLE
Understanding optimal military decision making: Year 2 progress report

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Quinn Kennedy, Peter Nesbitt, Jon Alt, Ron Fricker

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Naval Postgraduate School (NPS-Monterey) 14,973.00
1 University Circle

Monterey, CA 93943 -5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
62626-MA.4

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
This research aims to gain insight into optimal wargaming, decision-making mechanisms using neurophysiological measures by investigating whether brain activation and visual scan patterns predict attention, perception, and/or decision-making errors through human-in-the-loop wargaming simulation experiments. We report preliminary results from a study in which 34 military officers completed military-relevant tasks that tap into reinforcement learning and cognitive flexibility, while their eye gaze and brain activity was monitored via eye-tracking and electroencephalography (EEG) technology. Results indicated that the tasks successfully elicited reinforcement

15. SUBJECT TERMS
Military decision making, wargames, neurophysiological predictors

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Quinn Kennedy
					19b. TELEPHONE NUMBER
					831-656-2618

Report Title

Understanding optimal military decision making: Year 2 progress report

ABSTRACT

This research aims to gain insight into optimal wargaming, decision-making mechanisms using neurophysiological measures by investigating whether brain activation and visual scan patterns predict attention, perception, and/or decision-making errors through human-in-the-loop wargaming simulation experiments. We report preliminary results from a study in which 34 military officers completed military-relevant tasks that tap into reinforcement learning and cognitive flexibility, while their eye gaze and brain activity was monitored via eye-tracking and electroencephalography (EEG) technology. Results indicated that the tasks successfully elicited reinforcement learning and cognitive flexibility, and that a suitable range of variability in performance occurred. Preliminary results of eye tracking provided insight into which pieces of information the subjects used in making their decisions. Several statistical methods for modeling the transition from naive decision making to experienced decision making are examined.

UNDERSTANDING OPTIMAL MILITARY DECISION MAKING

YEAR 2 PROGRESS REPORT

Quinn Kennedy, Ph.D.[1]; MAJ Peter Nesbitt,[2] USA; LTC Jon Alt,[1] USA; &
Ron Fricker, Ph.D.[1]

[1]Operations Research Department, Naval Postgraduate School, Monterey, CA

[2]TRADOC Analysis Center, Monterey, CA

ABSTRACT

This research aims to gain insight into optimal wargaming, decision-making mechanisms using neurophysiological measures by investigating whether brain activation and visual scan patterns predict attention, perception, and/or decision-making errors through human-in-the-loop wargaming simulation experiments. We report preliminary results from a study in which 34 military officers completed military-relevant tasks that tap into reinforcement learning and cognitive flexibility, while their eye gaze and brain activity was monitored via eye-tracking and electroencephalography (EEG) technology. Results indicated that the tasks successfully elicited reinforcement learning and cognitive flexibility, and that a suitable range of variability in performance occurred. Preliminary results of eye tracking provided insight into which pieces of information the subjects used in making their decisions. Several statistical methods for modeling the transition from naive decision making to experienced decision making are examined.

EXECUTIVE SUMMARY

MOTIVATION

As the Army focuses on enhancing leader development and decision making to improve the effectiveness of combat forces, the importance of understanding how to effectively train decision makers and how experienced decision makers arrive at optimal or near-optimal decisions has increased. Currently, there is little understanding of how military decision makers arrive at optimal decisions and the measurement of decision-making performance lacks objectivity. The use of neurophysiological measures in human-in-the-loop wargames has the potential to fill this knowledge gap and provide more objective measures of decision-making performance.

PURPOSE

This project's purpose is to investigate the role between neurophysiological indicators and optimal decision making in the context of military scenarios, as represented in human-in-the-loop, wargaming simulation experiments. In this second-year effort, we focused on the development of optimal decision making when all subjects begin as naïve decision makers. Specifically, we attempted to identify the transition from exploring the environment as a naïve decision maker to exploiting the environment as an experienced decision maker, via statistical and neurological measures.

ARMY RELEVANCY AND MILITARY APPLICATION AREAS

Objectively defining, measuring, and developing a means to assess military optimal decision making has the potential to enhance training and refine procedures supporting more efficient learning and task accomplishment. Through the application of these statistical and neurophysiological models, we endeavor to further neuromathematics and the understanding and modeling of decision-making processes to more deeply understand the fundamentals of Soldier cognition. This project supports the Army Training and Doctrine Command (TRADOC) Analysis Center's (TRAC's) fiscal year (FY) 14 research requirements: 1.2 – Agile Wargames, 2.6 – Mission Command Processes and Decision Making, and 2.2 – Enhancing Subject Matter

Expert (SME) Elicitation Techniques. The Veterans Affairs' (VA's) War-Related Illness and Injury Study Center (WRIISC) is interested in this project to help identify posttraumatic stress disorder (PTSD) and traumatic brain injury (TBI). The results of this project are also of potential interest to the Neurophysiology Office and Simulations Office in the Army Research Laboratories (ARL).

SUMMARY OF CURRENT STATUS

We developed two wargames and conducted a study that demonstrated that the wargames successfully elicit cognitive flexibility and reinforcement learning. Preliminary results will be reported at the 2014 Human Factors and Ergonomics Society Annual Meeting. We have merged and synchronized the decision, eye tracking, and EEG data for each subject. We are investigating several statistical methods to objectively define and assess the transition to optimal decision making, such as regret and sequential detection methods.

ACKNOWLEDGEMENTS

- The subjects, for volunteering their time, for their participation, and their feedback.
- Lieutenant Lee Sciarini, United States Navy, Ph.D., Assistant Professor at the Naval Postgraduate School (NPS), for demonstrating successful EEG calibration techniques and sharing his overall experience with the Advanced Brain Monitoring (ABM) EEG system.
- The Operations Research Department and Modeling, Virtual Environments and Simulation (MOVES) Institute for providing resources to purchase the necessary equipment and software.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

OVERVIEW

As the U.S. Army focuses on enhancing leader development and decision making to improve the effectiveness of combat forces, the importance of understanding how to effectively train decision-makers and how experienced decision-makers arrive at optimal or near-optimal decisions has increased (Lopez, 2011). Two cognitive characteristics necessary for military personnel to reach optimal decision making are reinforcement learning, the ability to learn from trial and error; and cognitive flexibility, the ability to recognize when the rules have changed or that the current strategy no longer works (Vartanian & Mandel, 2011). Although many laboratory tests of reinforcement learning and cognitive flexibility exist, these tasks may not necessarily capture military decision making due to the high stakes and uncertain environment in which military decisions are made. Assessment tools that leverage wargames (i.e., simulations of realistic military scenarios) to evaluate these two cognitive characteristics are needed. We determined that two common psychological tests that measure reinforcement learning and cognitive flexibility, the Iowa Gambling Task (IGT) (Bechara, Damasio, Damasio, & Anderson, 1994) and the Wisconsin Card Sorting Task (WCST) (Grant & Berg, 1948) could be modified to provide a more realistic military context as a first step towards understanding military decision making. (For an in-depth review of decision making, the IGT, and the WCST, see (Nesbitt et al., 2014).

The IGT was developed to measure prefrontal damage (Bechara et al., 1994). Persons with prefrontal damage tend to have difficulty detecting the long-term consequences of their decisions and actions. In this task, subjects receive a loan of $2,000 of play money and are asked to make a series of decisions to maximize the profit on the loan. Each decision entails selecting one card at a time from any of four available decks of cards (decks A-D). All cards give money and some cards also issue a penalty. Decks differ in the amount of money given on a single trial ($50 or $100), as well as the frequency and severity of penalties ($0 to $1,250). Healthy subjects should learn through reinforcement learning which decks have the best long-term payoffs (decks C and D) (Bechara, Damasio, & S.W., 1994; Steingroever, Wetzels, Horstmann, Neumann, & Wagenmakers, 2013). Main measures of decision performance are total money won and an advantageous selection bias (the proportion of good decks selected minus the proportion of bad decks selected).

The WCST taps the working memory, shifting, and inhibition components of executive function (Grant & Berg, 1948). Subjects view five cards, one card displayed at the top center of the screen, the remaining four displayed across the bottom of the screen. Each card contains symbols that vary in number, shape, and color. Over several trials, subjects try to figure out the matching rule that will correctly match the card on the top of the screen with one of the four cards at the bottom of the screen. Unbeknown to the subjects, the matching rule changes once they have 10 consecutive correct matches. For example, after 10 consecutive correct matches based on the color of the symbols, the matching rule could then change to the number or shape of the symbols. Thus, subjects must not only learn and maintain in working memory the correct matching rule while inhibiting irrelevant stimuli, but also exhibit cognitive flexibility in detecting when the rule has changed (Grant & Berg, 1948). The task is complete when subjects either successfully complete two rounds of each matching rule, or 128 trials. Main performance measures include total percentage correct, percentage of perseverative responses (the number of incorrect responses that would have been correct for the previous matching rule), the number of matching rules achieved, and the total number of trials completed (fewer indicates better performance).

The purpose of this study was to first modify two existing cognitive assessments that measured reinforcement learning and cognitive flexibility in order to assess active duty military officers' decision-making behavior on these tasks. The convoy task, in which subjects incur or receive enemy or friendly damage, is analogous to the IGT, whereas the map task is modified from the WCST. In order to gain further insight into how military decision makers value information, eye-tracking data was captured for each subject during each task. Numerous studies indicate that eye-movement data via eye-tracking technology can provide valuable insights into subjects' attention allocation patterns and underlying cognitive strategies during real-world tasks (Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001; Marshall, 2007; Sullivan, Yang, Day, & Kennedy, 2011). To assess whether the convoy and map tasks successfully elicit reinforcement learning and/or cognitive flexibility, we tested the following predictions:

(1) Convoy Task: Subjects will demonstrate reinforcement learning by having a positive advantageous selection bias, and by correctly reporting which routes are the safest and the most dangerous.

(2) Map Task: Subjects will demonstrate cognitive flexibility by having low rates of perseverative responses, completing at least three matching rules, and having at least 70% correct trials.

(3) Exploratory analyses from the eye-tracking data will provide insights into subjects' prioritization of information.

The second purpose of this study is to begin to statistically model the transition from naïve decision making, in which exploration of the options occurs, to experienced decision making, in which exploitation of the options takes place. Nesbitt et al. (2014) provide an overview of several possible methods; in this report, we focus on regret and sequential detection methods that use trial-by-trial latencies to detect exploration-exploitation mode changes: the exponentially weighted moving average of latencies and the sequential sample variances of latencies.

REGRET

Regret is the difference of a participant's single trial outcome and the outcome from the ideal decision, given perfect knowledge. Less regret is better; on any given trial, regret can be zero if the participant selects the best decision. More generally, absolute regret compares the outcome of participant actions to the outcome generated by playing the optimal policy at each of the *n* trials. Given K ≥ 2 routes and sequences $ri_{,1}$, $ri_{,2}...r_{i,n}$ of unknown outcomes associated with each route i = 1,...K, at each trial, t = 1,...n, participants select a route I_t and receive the associated outcomes $r_{It,t}$. Let $r^*_{i,t}$ $s_i^2 > h$ be the best possible outcome possible from route i on trial t; (Auer & Ortner, 2010). The regret after n plays $I_1,...I_n$ is defined by

$$R_n = \sum_{t+1}^{n} r^*_{i,t} - \sum_{t+1}^{n} r_{I,t} .$$

Regret provides insights in the aggregate over the course of a set of *n* trials (i.e., total regret) and, when examined, per trial. Regret per trial provides a measure of a participant's ability to identify the best choice available at a given point in time.

THE EXPONENTIALLY WEIGHTED MOVING AVERAGE OF LATENCIES

Let us start with the former, where we could use the exponentially weighted moving average (EWMA) method drawn from the statistical process control literature (Fricker, 2010). Let x_i denote the latency at time i, $i = 2, 3, \ldots, 100$ (where, presumably, there is no latency at time $i = 1$). Then, at time i, we would monitor

$$E_i = \alpha x_i + (1 - \alpha)E_{i-1},$$

where α is a smoothing parameter, $0 < \alpha \leq 1$, and, typically, the method starts by setting $E_1 = x_2$. Here, we assume that at time $i = 1$ the subject starts out in the exploration mode and the question is to identify when he or she switches to exploitation. This is done by setting a threshold h and the first time i that $E_i < h$ we declare that the subject is now in exploitation mode.

Three questions then arise: (1) how to choose α? (2) how to choose h? and (3) is h subject specific?

MONITORING SEQUENTIAL SAMPLE VARIANCES

Given the questions that need to be addressed in using the Exponentially Weighted Moving Average of latencies, monitoring latency variance may be easier to implement than monitoring the mean since, when a subject goes into exploitation mode, it is possible that the variance will get close to zero (for all subjects). This method is one way to implement a sequential scheme, where we would monitor the sample variance calculated from moving windows of data. Specifically, as before, let x_i denote the latency at time i, $i = 2, 3, \ldots, 100$. Then, for some window of data of size $w + 1$, starting at time $i = w + 2$, sequentially calculate

$$s_i^2 = \frac{1}{w} \sum_{j=i-w}^{i} \left(x_j - \bar{x}_i \right)^2,$$

where

$$\bar{x}_i = \frac{1}{w+1} \sum_{j=i-w}^{i} x_j .$$

The idea is to monitor $s_{w+2,}^2 s_{w+3,}^2 s_{w+4,}^2 \ldots$ and when it is less than some threshold h, we declare that the subject has gone from exploration to exploitation.

For this method, the question is how to choose w. There are two considerations: (1) $w + 1$ should be smaller than the smallest length of time a subject might be in exploration mode when the experiment first starts, and (2) smaller is better in the sense that the method will more quickly indicate the shift to exploitation, but w+1 cannot be so small that the sample standard deviation estimates are too variable because of excess noise. Ultimately, we will want to do some simulations to see what a good choice for w might be. Our initial guess would be something in the range $0.5 \leq w \leq 5$ or so.

Now, there is also the question of how to detect whether someone reverts from exploitation back to exploration. One possibility would be to continue to monitor the sample variances and, once someone is in exploration mode, should $s_i^2 > h$, then we say they have reverted back to exploration. However, it may be that we need two thresholds, call them h_1 and h_2, where $h_2 > h_1$, which would work as follows. For someone in exploration mode, then they only switch to exploitation at time i when $s_i^2 < h_1$, while for someone in exploitation mode, they only switch to exploration at time i when $s_i^2 > h_1$. The key idea here is that having two thresholds with some separation between them may decrease inadvertent (i.e., excessive) switching back and forth between modes due to noise in the data.

METHODS

SUBJECTS

The study collected data from 34 military officers from all branches of service: 9 U.S. Army, 11 U.S. Marine Corps, 10 U.S. Navy, 3 U.S. Coast Guard, and 1 U.S. Air Force. The mean age was 35.11 years (standard deviation (SD) 4.9) with a mean time in service of 12.7 years (SD 4.42), of which the average time deployed was 19.57 months (SD 12.12) (note that one subject did not report their deployment time). Of the 31 subjects with deployment experience, the mean time since their last deployment was 37.98 months (SD 25.18) and 19 of those deployments were to ground combat zones (Iraq or Afghanistan). A majority of the subjects (n=24), served as staff officers during their most recent deployment. The majority of the subjects were male (30 males, 4 females) and the majority of subjects possessed 20/20 or better visual acuity (n=29). Subjects were recruited through bulk email to all NPS students, faculty, and staff; posting of flyers; and word of mouth.

DECISION-MAKING TASKS

Two decision-making tests were administered: the convoy task and map task.

Convoy Task: Our version of the IGT, the convoy task, serves as a simple wargame. In the convoy task, subjects are asked to select one of four possible routes, over an unknown number of trials, to maximize the damage to enemy forces, while minimizing the friendly damage accrued over all trials. These routes are analogous to the decks of the original IGT. At each trial, the subject is provided immediate feedback in the form of three separate pieces of information: a reward, a penalty, and a running total. The reward, the number of enemy forces damaged, is called *Damage to Enemy Forces*. The penalty, the number of friendly forces damaged, is called *Damage to Friendly Forces*. The running total is called *Total Damage*, defined as the previous trial's value of Total Damage plus the previous trial's Damage to Enemy Forces minus the previous trial's Damage to Friendly Forces. The units of value are in damage. Damage to Enemy Forces is considered positive in value (damage given to the enemy) and desirable to the participant. Damage to Friendly Forces is negative in value (value lost due to damage to friendly forces) and is not desired by the participant. The subject seeks to determine which route to select on the next turn through repeated sampling of routes. A participant selects routes until the end, not knowing that the task will complete after 200 selections. The assumption is that the subject maintains some estimate of the value similar to Accumulated Damage for each route and updates the estimate after each trial. The accuracy of the estimate will vary between subjects, as will the manner in which the subjects incorporate information indexed by trial into their estimate.

The feedback for the convoy task is derived from the first published IGT. The convoy task payout schedule for each route demonstrated in Appendix A is constructed from the original IGT schedule. Each route has its own "deck,"—a scripted, ordered set of specified values. For example, every participant will find that the third time they pick route *A*, it returns +100 and –150. Even though these returns by route are set and are the same for each participant, the games will progress differently due to the divergence of route selection between participants. Table 1 provides summary statistics of the returns for each route. The convoy task offers minimal visual difference between images representing the available options (see Figure 1). The intent of similar-looking options is to minimize the visual bias, an intent consistent with the first IGT (Bechara, Damasio, Tranel, & Damasio, 2005).

Route A		**Route B**		**Route C**		**Route D**	
Min.	–250	Min.	–1,250	Min.	0	Min.	–200
25%	–150	25%	100	25%	0	25%	50
Median	25	Median	100	Median	25	Median	50
Mean	–25	Mean	–25	Mean	25	Mean	25
75%	100	75%	100	75%	50	75%	50
Max.	100	Max.	100	Max.	50	Max.	50

Table 1. Summary statistics for the damage that can occur for each route during the convoy task. Negative numbers indicate friendly damage; positive numbers indicate enemy damage.

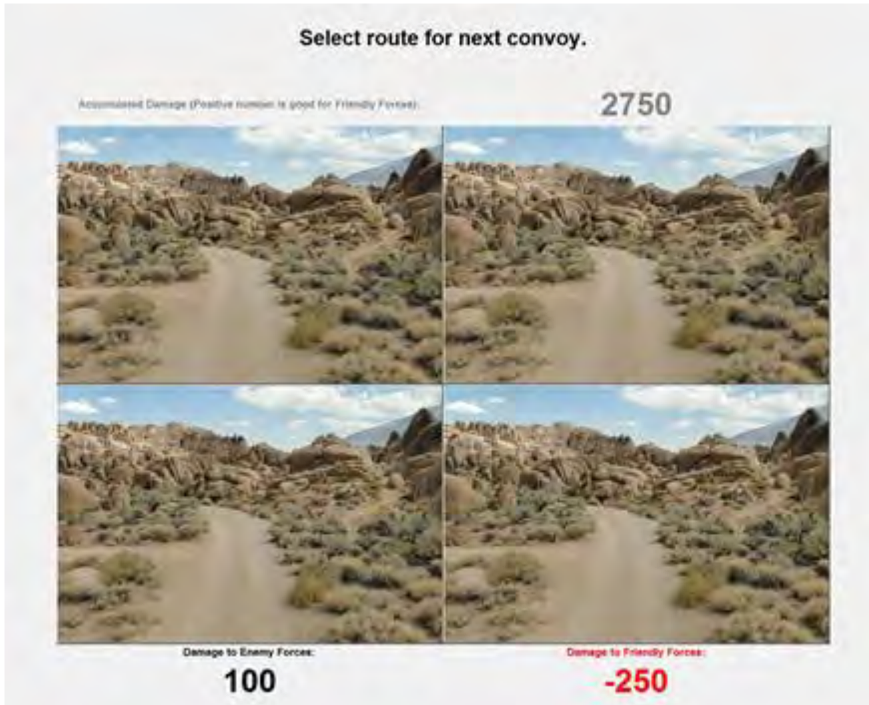


Figure 1. Screen shot of the convoy task in piloting; a typical subject's view of the task. We see that the participant's last choice caused 100 damage to the enemy (Damage To Enemy Forces) and a loss of –250 to friendly forces (Damage to Friendly Forces) resulting in a trial loss of –150 (not shown). The Accumulated Damage is 2,750. A positive Accumulated Damage value is desirable to the participant. Notice that four routes are represented by the same image.

CONVOY TASK MEASURES

- ***Total Damage*****:** All subjects start with 2,000 enemy damage. Therefore, the Total Damage is calculated as the difference between the initial Damage Score and the last Damage Score at the end of 200 trials. Total damage significantly larger than 2,000 demonstrates optimal decision performance, whereas total damage at or below 2,000 indicates suboptimal decision performance.
- ***Frequency of Friendly Damage*****:** The number of trials in which friendly damage occurred.
- ***Frequency of Heavy Friendly Damage*****:** The number of trials in which friendly damage of –1,250 occurred, which is the highest amount of friendly damage that can occur.
- ***Advantageous Selection Bias*****:** The typical decision performance measure from the IGT is the advantageous selection bias, in which the proportion of bad routes

selected is subtracted from the proportion of good roads selected. According to the IGT, routes 3 and 4 are considered good; 1 and 2 are considered bad. Positive advantageous selection bias scores indicate a propensity to select the good routes, whereas negative scores indicate a tendency to select the bad routes.

- ***Route Selection*:** Route selection is the frequency with which the subject selected each route over all trials.
- ***Trial Latency*:** Latency is defined as the amount of time that subjects take to make a decision on each trial. It is measured as the amount of time taken between key press selections from trial to trial.

MAP TASK

Our military-relevant version of the WCST is the map task. In the map task, subjects view five maps, with one map displayed at the top center of the screen and the remaining four displayed across the bottom of the screen. Figure 2 is a typical subject's view of the task. The maps are analogous to the cards of the original WCST. Each map contains military graphic control graphics that vary in meaning, color, and shape. These graphics are described in Figure 3 and developed from U.S. Army FM 1-02, *Operational Terms and Graphics* (United States Army, 2004). Subjects are asked to match one of four lower maps to the top one over an unknown number of trials.

Figure 2. Screen shot of the map task in piloting; a typical subject's view of the task. On this trial, the subject should sort on intended action graphics (black) and, therefore, should select the map on the far right.

	friendly graphics	intent graphics	enemy graphics
Level 0	no friendly graphic	no intent graphic	no enemy graphic
Level 1	friendly armor platoon	ambush	enemy infantry squad
Level 2	friendly aerial vehicle	clear	enemy anti-armor squad
Level 3	friendly infantry platoon	block	enemy anti-air squad

Figure 3. Description of the graphics in the map task. There are three categories of graphics: friendly (colored blue), intent (colored black), and enemy (colored red). The sorting rules correspond to the same categories. Each category has four levels, each with a particular corresponding graphic.

Over several trials, subjects try to figure out the matching rule that will correctly match the map on the top of the screen with one of the four maps at the bottom of the screen. This process of matching maps is similar to card matching in the original WCST; unknown to the subject, the matching rule changes once the subject has 10 consecutive correct matches. For example, after 10 consecutive correct matches by sorting the maps using the sorting rule based on the friendly graphic, the matching rule changes to sorting maps according to the intent graphic. The task is completed when either the subject has successfully completed two rounds of each matching rule or until they have completed 128 trials.

Map Task Measures

For the map task, we use the same decision performance measures developed from WCST.

- ***Number of Trials*****:** Total number of trials taken to achieve all six sorting rules or the subject has reached the maximum of 128 trials.
- ***Total Percent Correct*****:** Number of trials in which the subject made the correct decision, divided by the total number of trials completed.
- ***Perseverative Responses*****:** The number of incorrect responses that would have been correct for the preceding category/rule.
- ***Perseverative Errors*****:** The number of errors in which the subject has used the same rule for their choice as their previous choice.
- ***Percent Perseverative Errors*****:** The number of perseverative errors, divided by the total number of trials.
- ***Nonperseverative Errors*****:** After excluding the perseverative errors, the number of other errors.
- ***Number of Trials to Complete First Rule*****:** Total number of trials needed to achieve the first 10 consecutive correct choices.
- ***Number of Rules Achieved*****:** The number of trials of 10 consecutive correct choices.
- ***Failure to Maintain Set*****:** The number of trials in which five or more consecutive correct choices occur without completing the category (i.e., without reaching 10 consecutive correct choices).

- ***Trial Latency*:** Trial latency is measured as the amount of time taken between key press selections from trial to trial.

SURVEYS

A demographics survey and posttask survey were used to quantify and categorize blocking factors, such as elements of military experience, and to collect qualitative responses from the subjects at the conclusion of the tasks.

Demographic Survey

The demographic survey in Appendix B was administered prior to the decision-making tasks. The survey includes questions regarding subjects' deployment history, as well as general demographic information such as age and rank.

Posttask Survey

The posttask survey in Appendix C was administered after the completion of the decision-making tasks. Subjects provided qualitative responses regarding their strategies for each decision-making task.

COVARIATE MEASURES

Because the decision-making tasks place demands on working memory and visual processing speed, we are including covariate measures of these cognitive functions. The tasks are also highly visual; therefore, a visual acuity test also is administered.

Digit Span Memory Test

The digit span forwards and backwards test measures working memory (Wechsler, 2008). In digit span forwards, the experimenter states a series of digits, starting with two digits, and the subject must repeat them back. The number of digits increases, with two trials per number of digits. The test is discontinued if the subject has an incorrect response to both trials for a particular number of digits. In digit span backwards, the same procedure is followed, except this time the subject must repeat the digits in the reverse order.

Trails A and B

Trails A and B test visual processing speed (Wechsler, 2008). In Trails A, the numbers 1 through 25 are randomly distributed on a paper. The subject starts at 1 and must draw a line to each number in chronological order. Subjects are instructed to work as quickly and accurately as they can. In Trails B, subjects now see both numbers and letters, and must connect 1 to A, A to 2, 2 to B, and so on until they reach Z. They also are instructed to work as quickly and accurately as they can.

Snellen Test

Because the decision tasks are visually based, the Snellen eye chart was used to measure subjects' visual acuity at the beginning of the experiment. The Snellen eye chart is placed on the wall and consists of 11 lines of block letters, in which each line of letters gets progressively smaller. Subjects stood 20 feet from the chart, cover one eye, and read aloud as many lines as they can. They then covered the other eye and read aloud as many lines as they could. The last line that the subject could accurately read for each eye is recorded.

EYE-TRACKING MEASURE

In this initial report, we used percentage dwell time as the main measure of eye tracking. Percentage dwell time is the percentage of time that the subject's eye gaze looked at a particular region of interest. For example, the percentage of time that a subject looked at their friendly damage score.

EEG MEASURES

The EEG software automatically provides real-time measures of distraction, sleepiness, engagement, and cognitive workload.

EQUIPMENT

The devices used in this study consisted of a laptop computer, two eye-tracking stereo cameras, a desktop computer, and an EEG. The laptop runs FaceLAB 5.0.7 software on a Windows XP operating system. The stereo cameras supply data to FaceLAB on the laptop. FaceLAB software and the stereo cameras were made by Seeing Machines, Inc. The desktop computer runs the EyeWorks data collection suite and ABM Visual software on the Windows 7 operating system. The laptop has a 15-inch screen that is not viewed by the subjects. The desktop uses a 30-inch primary monitor that is viewed by the subjects, and a 24-inch secondary monitor that is not viewed by the subjects.

The stereo cameras use 12 millimeter (mm) lenses to detect infrared light reflected off the subjects' eyes and face to monitor the position of the head and direction of the eye gaze. These data are fed from the laptop to the EyeWorks Record software on the desktop.

EEG data is recorded through an ABM X10 B-Alert Headset through nine channels (F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4) and sent through a wireless connection to B-Alert Visual software on the desktop.

Other materials used include 70% ethyl alcohol to clean the subjects' mastoid reference points, Synapse brand electrolytic gel, and recording electrodes provided by ABM.

PROCEDURES

The subjects completed the experiment in a single visit. Upon arriving at the test location, they first completed the IRB-approved consent form, followed by the demographic survey, and the cognitive tasks, including the digit span forward/backward task and two forms of the trail-making test. Next, the Snellen visual acuity test was completed. The next step entailed EEG and eye-tracking calibration. Eye-tracking calibration includes verifying the integrity of the camera configuration, building a personalized head model for the subject, and calibrating the subject's gaze with respect to the screen. EEG calibrating tasks include getting scalp and reference impedance levels under 40 kOhms and creating a baseline EEG profile using the three-choice vigilance, eyes open, and eyes closed tasks. Once all calibration steps are satisfied, the subject completed the convoy task, followed by the map task. Finally, they provided their responses to the posttask survey.

DATA MERGING AND SYNCHRONIZATION

The EEG and decision data were matched using the system time from the raw EEG data files and the system time from the decision data files as a key. During data collection, a marker was placed in the EEG data to identify when each subject actually began the convoy task. The system times corresponding to these markers were manually collected postexperiment and used to identify the true start point in the data for each subject. Raw EEG observations were matched to a behavioral trial if their system time was greater than or equal to the start time for the trial, and less than or equal to the start time of the subsequent trial. This mapping of raw EEG to behavioral trials was then used to map the processed EEG data provided by the EEG software which is aggregated based on "epochs." Each observation in the raw data file is assigned an epoch and this corresponds to the epochs used in the processed EEG data. Epochs were matched to trials based on the behavioral map in the previous processing step.

INITIAL RESULTS

We first provide initial decision-making results from the convoy and map tasks, along with results that investigated any relationships between military demographics, covariate measures, and decision performance measures. Next, preliminary results from the eye tracking and EEG are presented. Finally, results exploring sequential detection methods in modeling the transition from exploration to exploitation are described.

CONVOY TASK RESULTS

Decision Results

All analyses utilized a two-tailed 0.05 alpha level. Although mean total damage score was above 2,000 and the advantageous selection bias was positive, results were not significant (p's > 0.05) (see Table 2). As would be expected, the total damage score was negatively correlated with the number of high friendly damage, ($r = -0.87$, $p < 0.001$) and frequency of friendly damage ($r = 0.39$, $p < 0.05$), but very strongly positively associated with advantageous selection bias ($r = 0.97$, $p < 0.001$). Subjects also successfully distinguished between safe and dangerous roads, ($\chi 2$ (3) = 23.63, $p = 0.005$). In a question asking subjects to rank order the routes from safest to most dangerous, 42% reported route 4 as the safest, followed by route 3 (27%), whereas 42% of subjects reported route 1 as the most dangerous, followed by route 2 (33%). Table 2 reveals that subjects benefited from having 200 trials instead of 100. Results from paired t-tests indicated that the advantageous selection bias improved in trials 101-200 compared to trials 1-100 ($t(33) = 2.87$, $p = 0.007$), and a trend for people to learn to avoid high friendly damage ($t(33)= 1.85$, $p = 0.07$) in the second half of the wargame. Improvements in decision performance were due to the decrease in route 2 selection ($t(33)=2.70$, $p = 0.01$) and an increase in route 3 selection ($t(33) = 1.87$, $p = 0.07$). Improvements in decision performance over time are captured in Figure 4, which indicates that only after about trial 125 did subjects' total damage, on average, exceed the baseline of 2,000. Figure 4 also illustrates the large range of variability in decision performance.

Performance Variables	First 100 Trials Mean (sd)	Trials 101 - 200 Mean (sd)	All 200 Trials Mean (sd)
Total damage score # trial	2,077.94 (883.96)	N/A	2,402.94 (1,725.69)
Number of trials with friendly damage	24.50 (6.46)	26.65 (7.44)	51.15 (11.05)
Number of trials with heavy friendly damage	3.62 (1.39)	3.06 (1.72)	6.68 (2.59)
Route selection frequency (%)	13.82 (7.88)	12.56 (8.59)	13.19 (7.27)
Route 1	38.91 (14.30)	30.74(16.84)	34.82 (12.82)
Route 2	21.62 (16.59)	28.77 (20.63)	25.19 (15.02)
Route 3	25.64 (12.93)	27.94 (18.48)	26.79 (12.39)
Route 4	–5.47 (30.73)	13.41 (41.57)	7.94 (62.38)
Advantageous selection bias			
N/A = Not applicable; as it is not possible to calculate this particular variable.			

Table 2. Descriptive statistics of convoy task decision variables for the first 100 trials, trials 101-200, and all 200 trials.

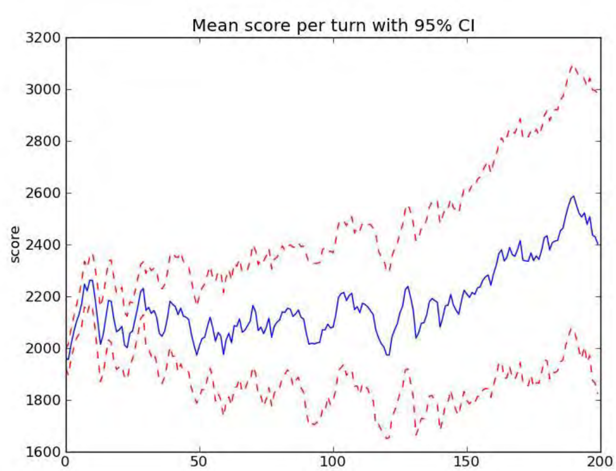


Figure 4. Mean total damage score per trial (blue line) with 95%CI (red dotted lines). Subjects begin with 2,000 total damage.

Next, exploratory analyses were conducted to determine if decision performance could be explained by cognitive function or demographic characteristics of the subjects. Surprisingly, Trail B time was positively associated with better decision performance; this association was driven by subjects' decision performance during the first 100 trials. Increased Trail B time was associated with high total damage score (r = 0.47, p = 0.006), and better advantageous selection bias (r = 0.347, p = 0.048), and fewer trials in which heavy friendly damage was incurred (r = –0.335, p = 0.057). A similar pattern is seen if Trail B normed data is used. Figure 5 illustrates this pattern. No other cognitive test or demographic characteristic (e.g., age, military rank, service branch) was associated with convoy task decision performance.

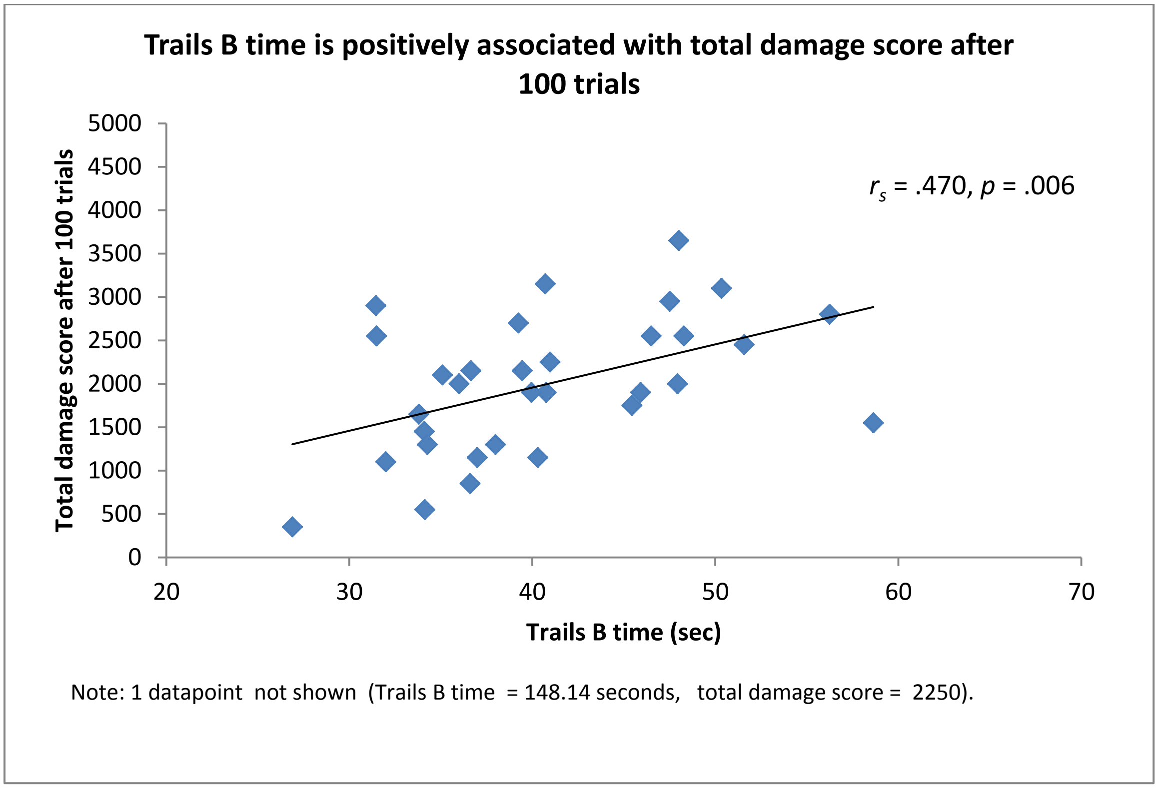


Figure 5. Longer time to complete Trails B is associated with higher total damage score at the end of 100 trials.

Latency Response Results

We created a latency response variable, which was calculated as the proportion of trials in which a subject's decision immediately after receiving feedback of moderate or heavy friendly damage was greater than 2 sd above their baseline time. Mean latency response to heavy friendly damage was 30% (sd = 23.5%) with a range from 0% to 100%. Mean latency response to moderate friendly damage was 18.2% (sd = 12%) with a range of 0% to 52.9%. There was a trend in the association between percentage of long latencies after heavy friendly damage and total damage score (r = 0.297, p = 0.093). Additionally, there was a positive correlation between the percentage of long latencies after medium friendly damage and total damage score (r = 0.380, p = 0.029). These results cannot be explained by subjects' processing speed or working memory, as neither latency response was associated with Trails A or B, or digit span forwards or backwards. Importantly, mean latency was not associated with either total damage score or advantageous selection bias.

REGRET

Regret indicates the difference between a participant's decision and the optimal decision, based on perfect knowledge of the payout schedule of each route. To provide an overall sense of participants' regret over the 200 trials, we first separated the participants into two groups by classic performance measures of final damage and the advantage selection bias using Ward Hierarchical Clustering, using euclidean distance. Clustering separates the sample cleanly into two groups: high and low performers. As illustrated in Figure 6, at about trial 50, the high performers' regret steadily decreases, indicating that their decisions over trials became steadily more optimal. In contrast, the low performers' regret remains high throughout the task.

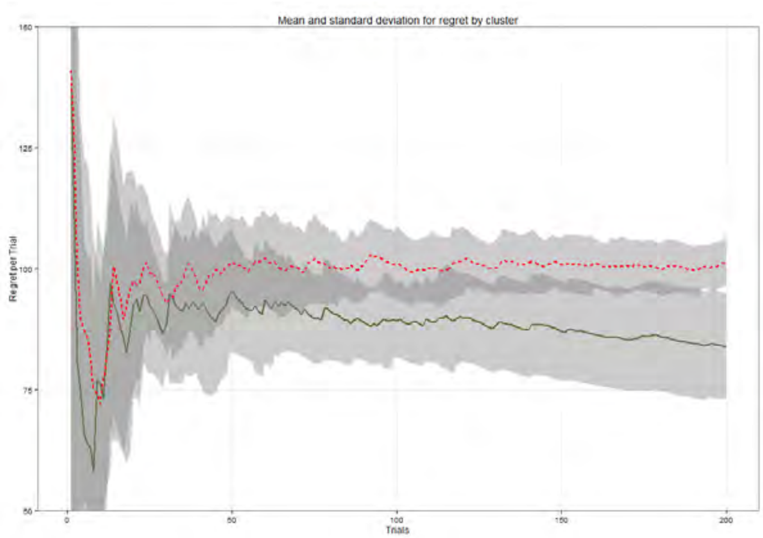
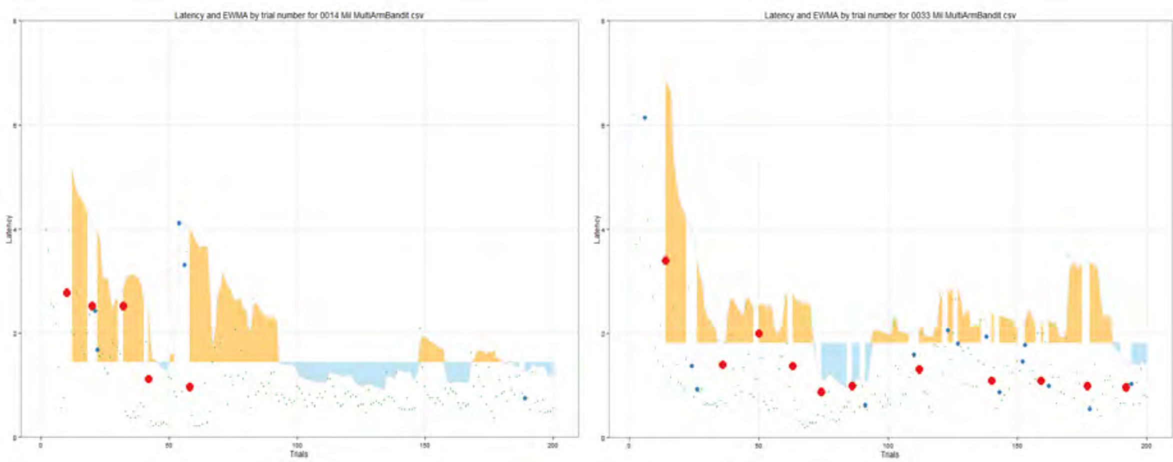


Figure 6. Cluster analysis revealed high-performing and low-performing groups based upon classic measures of IGT performance, total damage score, and advantageous selection bias. The high-performing group's regret per trial (solid green line) steadily drops after about 50 trials, whereas the lower-performing group (dashed red line) remains at approximately 100. The gray shading represents each confidence interval at one standard deviation, while the overlap is represented with dark gray. Regret per trial is a measure of the participant's ability to identify the best route available at a given point in time.

SEQUENTIAL DETECTION METHOD: USING LATENCY DATA TO DETERMINE EXPLORATION VS. EXPLOITATION COGNITIVE STATES

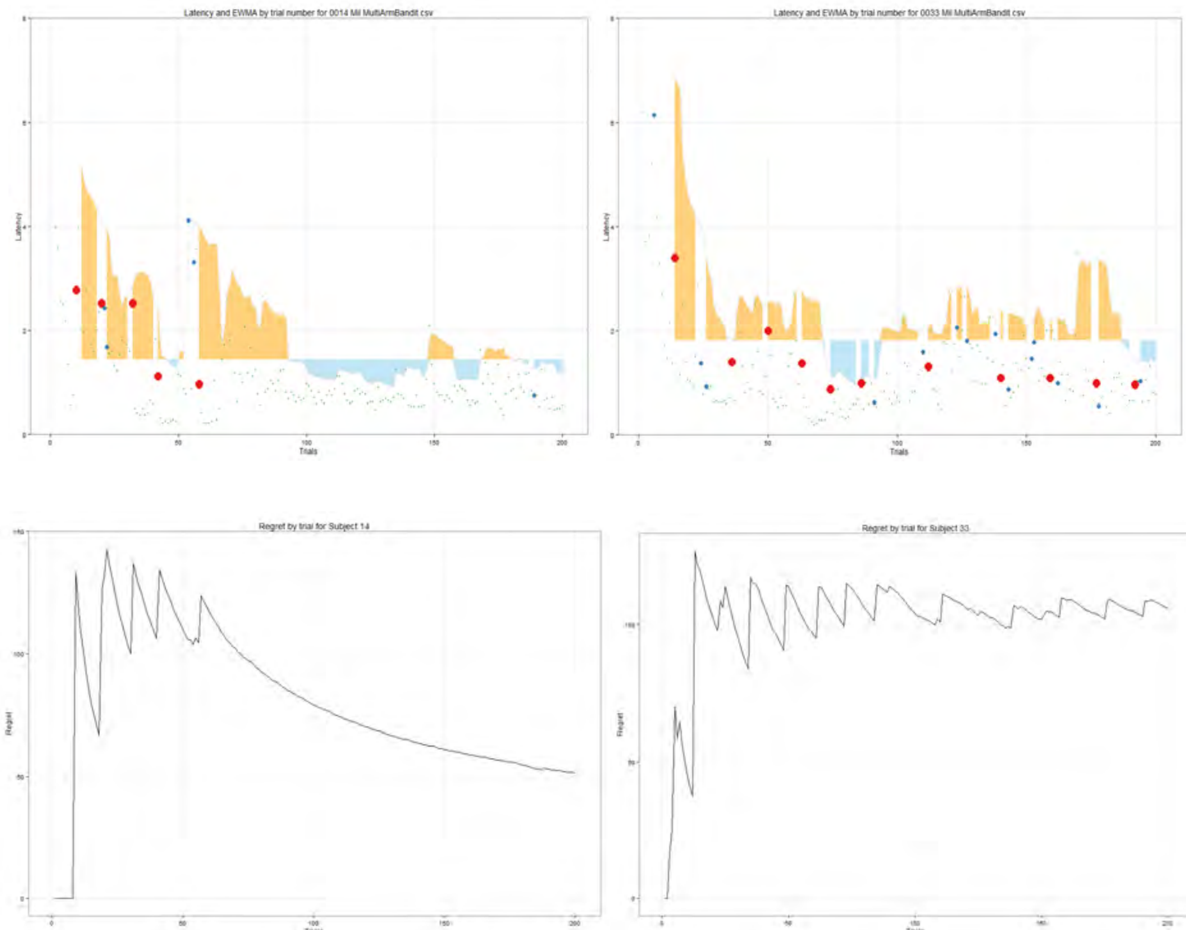
As illustrated in Figures 7a and 7b, we successfully used variability in trial-by-trial latency time to detect periods of exploration and exploitation cognitive states. A single explore/exploit latent threshold was developed for each subject, derived from twice the standard deviation above and below all latency times for 0 or 50 friendly damage (i.e., the baseline latency time) for that subject. Therefore, exploration was defined as trials in which the latency time was at least 2 SD higher than the baseline latency time. Exploitation was defined as two SD lower than the baseline latency time. Note that these definitions do not take into account actual decision performance, but solely the subject's cognitive state at a given time in the task. Figures 7a and 7b depict two distinct patterns of exploration and exploitation. Figure 7a depicts an optimal exploration to exploitation transition, whereas Figure 7b illustrates a pattern of primarily exploration throughout most of the task.



Figures 7a and 7b. Use of sequential sample variances in latency times to determine exploration and exploitation cognitive states.

COMBINING SEQUENTIAL DETECTION METHODS WITH REGRET

The combination of trial-by-trial information regarding the subject's current cognitive state (exploration or exploitation) with actual performance (measures of regret) provides insights into whose cognitive state is aligned with actual performance. In Figures 8a through 8d, we see that although subjects 14 and 33 show distinct differences in cognitive state, their cognitive state is aligned with their measure of regret. Subject 14 goes through a period of exploration until about trial 90, at which point they are predominantly in exploitation mode. Consistent with this cognitive state pattern, subject 14's regret is quite high until about trial 90, at which point it begins to steeply decrease. Recall that lower regret means that the subject's decisions are verging towards the best possible decision. Thus, when subject 14's cognitive state is in exploration mode, their regret is correspondingly high. When their cognitive state transitions to exploitation, their regret consistently decreases. In contrast, subject 33 maintains an exploration cognitive state throughout most of the task and, correspondingly, their regret is consistently high throughout the task.



Figures 8a-8d. Figures 8a and 8b show subject 14's and subject 33's exploration and exploitation cognitive states. Figures 8c and 8d depict the same subjects' regret, a measure of how much a subject's decisions deviate from the optimal decision over the course of the task. Figures 8a and 8c, and 8b and 8d, illustrate the concordant pattern between cognitive state and their actual decision performance as measured by regret for two different subjects.

Preliminary Eye-Tracking Results

Three subjects had unusable eye-tracking data; therefore, eye-tracking results are based upon 31 subjects. Preliminary eye-tracking analyses revealed that subjects spent most of the time looking at the routes and the least amount of time looking at the total damage score (see Table 3). Subjects relied more heavily upon friendly damage information than enemy or total damage. Subjects who tended to look at friendly damage also tended to look at enemy damage ($r = 0.442$, $p = 0.013$). There was a trend that the more subjects looked at the friendly damage, the higher was their advantageous selection bias ($r = 0.315$, $p = 0.08$).

Region of Interest (ROI)	Mean Percent (sd)
Total damage	5.49 (12.47)
Friendly damage	16.73 (14.87)
Enemy damage	6.55 (6.40)
Routes	71.23 (19.86)

Table 3. Mean number of fixations and percentage of time spent looking in each region of interest (ROI).

Preliminary EEG Results

As illustrated in Figure 9, the convoy task successfully elicited moderate levels of engagement and above average levels of cognitive workload. On average, low levels of distraction and very little sleepiness occurred during the task. Figure 9 also depicts the large amount of variability between subjects and between trials.

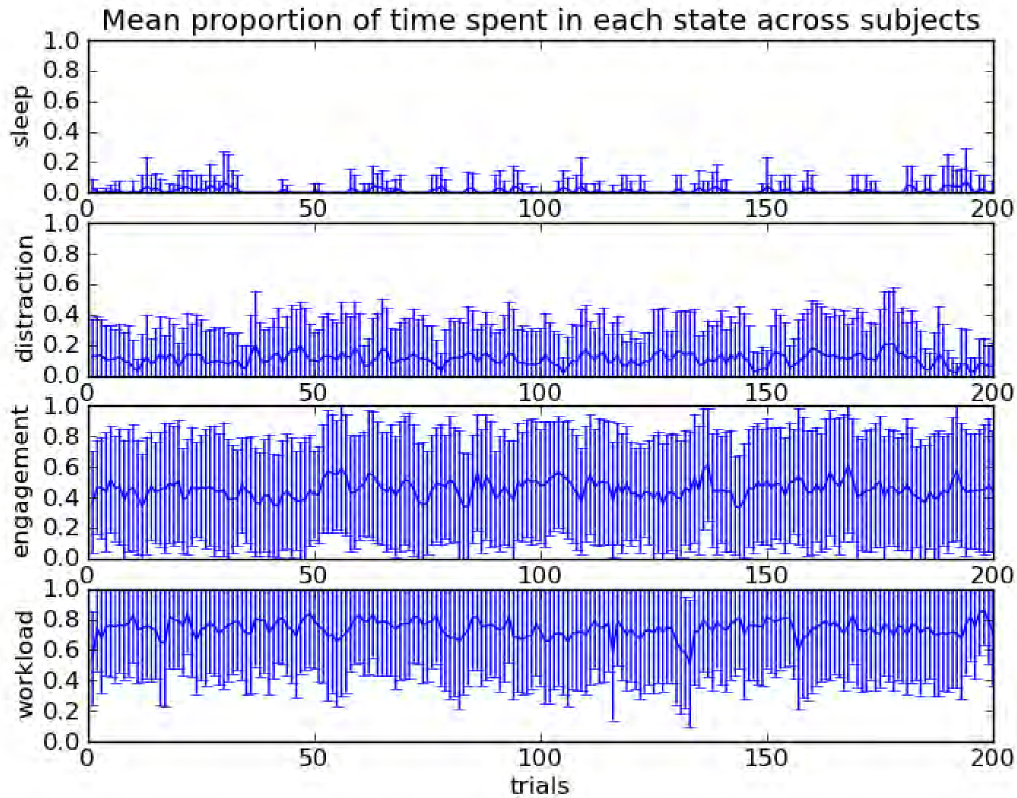


Figure 9. Mean proportion of time that subjects spent in a particular cognitive state across trials are indicated by the dark blue line. Error bars represent ± 1 sd.

Figure 10 illustrates the utility of combining neurophysiological and behavioral measures. Subject 33 had several periods of time when their workload level was high. Note that the peaks in latency time in the first several trials, and between approximately trials 160 to 170, overlap and/or precede peaks in cognitive workload. However, this subject was also frequently distracted and was minimally engaged in the task. Given insight into the subject's cognitive state throughout the task, it is not that surprising that subject 33 scored 700 in total damage, which was well below the average of 2,402.94.

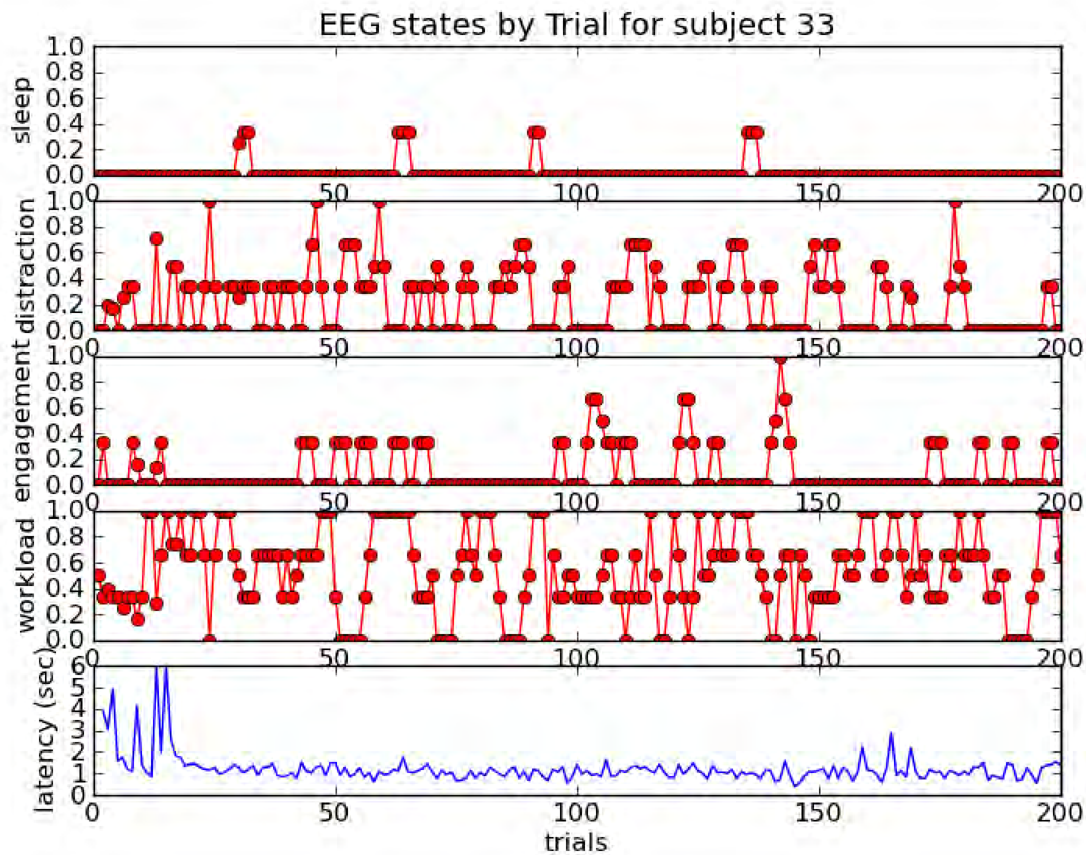


Figure 10. Illustration of pairing neurophysiological and behavioral measures of cognitive state. The bottom graph in blue represents subject 33's latency time on each trial. The graphs above shows the proportion of each trial that subject 33 spent being sleepy, distracted, engaged, or having cognitive workload.

Map Task Results

Results indicate that most subjects were able to determine the matching rules and that the matching rules changed periodically. Total percentage correct was not significantly different from 70% (95% CI: 59.81%-70.58%). Subjects completed an average of 3.21 matching rules (95% CI: 2.53-3.88). When subjects committed an error, they tended to be nonperseverative errors: On average, nonperseverative errors occurred on 33.56% (sd = 16.46%) of all trials, whereas perseverative errors occurred on 10% (sd = 8.79%) of all trials. Four subjects never completed the first matching rule. In the posttask questionnaire, 44% reported that they "immediately" recognized that the matching rule had changed, 29% "after a few trials," 15% "after several trials," and 12% "did not realize matching rule had changed." There was a positive correlation between how long it took subjects to realize that the matching rule had changed and the total number of trials completed ($r = 0.46$, $p < 0.05$), and a negative correlation between this self-reported variable and percentage of correct trials ($r = –0.53$, $p < 0.05$). As would be expected, longer mean latency was associated with needing more trials to complete the task ($r = 0.73$, $p = 0.0001$), making fewer correct decisions ($r = –0.72$, $p < 0.0001$), and fewer rules achieved ($r = –0.63$, $p < 0.0001$). Table 4 outlines subjects' performance on the main decision performance variables.

Variable	Mean (sd), Median, Range
Number of trials completed	119.35 (16.52), 128, 76-128
Percentage correct (%)	65.19 (15.43), 68.75, 36.72-86.25
Perseverative responses	11.82 (11.12), 9, 0-37
Nonperseverative errors	41.85 (22.52), 38, 8-81
Number of trials to complete first rule	42.9 (28.95), 34, 14-121
Number of rules achieved	3.21 (1.94), 4, 0-5
Failure to maintain set	2.32 (1.49), 2, 0-5

Table 4. Descriptive statistics of pilot subjects' performance on map task.

Eye-Tracking Result

Preliminary eye-tracking results indicate that the subjects spent the majority of their time looking at the example map at the top of the screen, and then appear to have spent more time looking at the cards in the center of the screen (maps 2 and 3), rather than maps on the farthest sides of the screen (maps 1 and 4).

ROI	Mean Percent of Time
Example map	46.95
map 1	6.12
map 2	14.00
map 3	21.77
map 4	11.16

Table 5. Mean percentage of time that subjects spent on each map.

DISCUSSION

Overall, results indicate that the modified tasks successfully captured reinforcement learning and cognitive flexibility. Results from the convoy task were consistent with other studies in which healthy adults completed the IGT (Steingroever et al., 2013). Although the total damage score and advantageous selection bias results were not significant, subjects correctly reported which routes were safe and which were dangerous. Subjects' scores on the modified IGT benefited from the additional 100 trials beyond the standard IGT protocol. Subjects' advantageous selection biases significantly increased due to a shift in route selection patterns, potentially attributable to the occurrence of reinforcement learning. Additionally, preliminary eye-tracking results indicate that subjects tended to prioritize information regarding friendly damage over information regarding total damage and enemy damage scores in making their decisions, highlighting the potential impact of the military context. Also consistent with previous studies of the IGT (Steingroever et al., 2013), all convoy measures showed large amounts of variability, suggesting that individual differences occur even among healthy subjects.

Importantly, objective measures of attention, latency response, and percentage of gaze spent in each region of interest, predicted decision performance on the convoy task. Latency response, the behavior of taking significantly longer to make a decision after receiving heavy or moderate friendly damage, provides an indicator of which subjects were actually paying attention to the feedback. The preliminary eye-tracking results indicate the underlying cognitive strategy subjects used in attempting to maximize their total damage score. Although subjects were instructed to maximize the total damage score, subjects rarely looked at the total damage score. Instead, of the three pieces of salient information (i.e., total damage score, enemy damage, and friendly damage), subjects focused primarily on friendly damage. Indeed, subjects who spent more time looking at friendly damage had higher total damage scores. Results from the map task were somewhat lower than what is typically found on the WCST for healthy subjects (Shan, Chen, Lee, & Su, 2008). However, subjects' perseverative response rates were relatively low, indicating that errors were not due to lack of cognitive flexibility. One reason that subjects may not have performed as well as predicted is because subjects' military experience actually may have made it harder for them to detect the matching rule. Unlike the original WCST, the symbols in the map task are meaningful. Each map can be "read" as a sentence by experienced military personnel: some type of friendly force should do an intended action upon an enemy force. Thus, these experienced military officers may have attempted to match the maps based upon meaning, rather than simply on color and shape. To date, we have focused solely upon the classical WCST measures in analyzing the map task data. Future goals include extending the successful statistical models already used to analyze the convoy task results to the map task, such as the use of latency response and change in latency variance. Additionally, analysis of the eye-tracking and EEG data will provide insight into participants' cognitive state during the task.

IMPLICATIONS OF INITIAL RESULTS

Combining real-time information regarding a participant's cognitive state as exploration or exploitation with actual decision performance has important training implications. First, it can be determined if the participant's cognitive state is aligned with their actual performance. As illustrated in Table 6, ideally, a participant is in the green cell in which they are in exploitation mode and their decision performance is optimal, as indicated by low regret. However, a participant's cognitive state also would be aligned if they are in exploration mode and their decision performance is nonoptimal (yellow cell). Ideally, a participant would begin in the yellow cell and transition to the green cell. When a participant's cognitive state is misaligned with actual decision performance, training intervention can occur (orange and red cell). Given that latency variance and regret can be measured in real time, the combination of these two measures can be used as a simple, near-immediate indicator of training intervention. Next, the incorporation of neurophysiological measures, such as eye tracking and EEG, can provide an understanding as to *why* a participant's cognitive state and actual performance are misaligned (see Figure 11). For example, perhaps a participant is in the red cell simply because they are not attending to the most relevant pieces of information. A participant in the orange cell may be experiencing an overly high cognitive workload during the task and therefore does not have the cognitive capacity to realize that they are performing well. Thus, these initial results suggest that highly efficient and target training interventions can occur with the combined use of decision performance, time to make a decision, eye-tracking, and EEG information monitored in real time.

		Cognitive State	
		Exploration	**Exploitation**
Decision Performance	High Regret	Seeking information and decision performance is not optimal	Acting upon acquired knowledge, but decision performance is not optimal
	Low Regret	Seeking information, yet decision performance is optimal	Acting upon acquired knowledge and decision performance is optimal

Table 6. Correspondence of exploration and exploitation cognitive states with actual decision performance, as measured by regret. Cell colors indicate the best (green) to worst (red) combinations of cognitive state and decision performance.

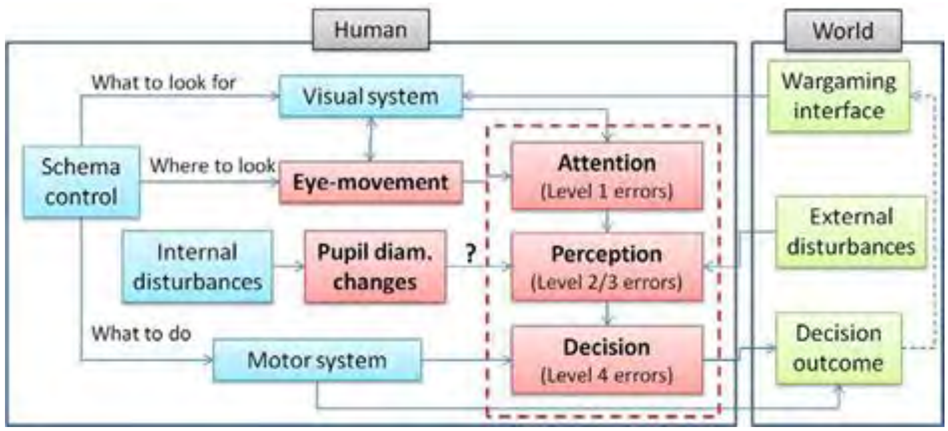


Figure 11. Model of nonoptimal decision making. Errors related to decision-making processes can be modeled in the following hierarchical levels. Level 1/attention errors occur if foveal vision (i.e., normal daylight vision) misses significant information. In this situation, it is obvious that optimal decision making cannot be reached. Level 2/perception errors occur when some important information is looked at, but not long enough for the human operator to perceive the information correctly. Level 3/perception errors occur when the human operator does not perceive the information due to internal/external disturbances. Level 2 and Level 3 errors can be distinguished via Bayesian modeling approach and EEG data. Finally, Level 4/decision errors can appear even when no attention or perception errors are associated. For example, decision outcomes can be nonoptimal due to inherent bias (e.g., the decision is preset by schema control, even before information has been scanned), within-subject differences, or between-subject differences.

SUMMARY

Wargames are a preferred method of training military personnel to make optimal military decisions. Wargames, however, are typically not assessed objectively and may not focus on training two cognitive functions necessary for optimal decision making: reinforcement learning and cognitive flexibility. The purpose of this study was to take the first steps to bridge the gap between the study of decision-making ability in the field of cognitive psychology and the study of decision making in a military setting. The use of well-known objective assessments to assess the effectiveness of training designed to improve reinforcement learning and cognitive flexibility shows great potential. Results demonstrate successful modification of the IGT and WCST into a military context. Future directions focus upon explaining individual differences in decision performance and using neurophysiological measures to identify why some participants performed well and others did not, as well as to more richly characterize exploration versus exploitation cognitive states. Future studies will examine military decision-making performance in sequential decision-making tasks with delayed rewards and more realistic military wargame scenarios.

CONCLUSION

FY14 PROGRESS

The following items generally list the measures of progress towards research project completion.

- *Study 1 conducted and completed.* Results from 34 subjects indicate that the wargames successfully elicit reinforcement learning and cognitive learning.
  - *Preliminary eye-tracking analyses*. Eye-tracking data revealed which information subjects used to make their decisions.
  - *Data merging and synchronization*. Successfully merged and synchronized decision and EEG data.
  - *Statistical methods.* Statistical methods to identify the transition from exploration to exploitation were implemented, such as sequential detection methods and regret.
  - *Eye-tacking consultation*. Dr. Ji Hyun Yang, an eye-tracking expert, worked with the team on cleaning the eye-tracking data and different ways to analyze this data.
  - *Journal articles and technical reports:*
    - Nesbitt, P., Kennedy, Q., Alt, J., Fricker, R., Whitaker, L., Yang, J., Appleget, J., Huston, J., & Patton, S. (2014). Understanding optimal decision-making in wargaming. Monterey, CA: Naval Postgraduate School. NPS-OR-14-001.
    - Nesbitt, P., Kennedy, Q., & Alt, J. Iowa Gambling Task modified for military domain. In submission to *Military Psychology*.
  - *Conference presentation*s:
    - Kennedy, Q., Nesbitt, P. & Alt, J. Assessment of cognitive components of decision making with military versions of the IGT and WCST. Accepted to the Human Factors and Ergonomics Society 2014 International Annual Meeting, October 27-31, Chicago, IL.
- *Student thesis study*, *completed.* Results based upon decision data indicate that tactical decision makers make the same decisions, in the same amount of time, with the same level of confidence in their decisions regardless of whether they

have a live or automated wingman. However, subjects with an automated wingman reported significantly lower trust in their wingman than subjects with a live wingman. See Appendix D.

- *Project meetings.* In the course of meeting objectives, the team met on a weekly basis; consultants joined the meeting on a monthly basis.
- *Equipment software procurement*. With funding from the Operations Research Department and MOVES Institute, NPS, we were able to purchase necessary equipment and software, including new computers, new eye-tracking computers, updated software and licenses, and a printer.
- *Collaboration with the War-Related Illness and Injury Study Center (WRIISC)*. WRIISC at the Veteran's Administration (VA) in Palo Alto, California has requested use of the convoy and map tasks to include in their battery of tests used to determine the cognitive functioning level of traumatic brain injury (TBI) patients. As WRIISC is one of the main VA research centers of TBI, the potential for productive collaboration is great.

INITIAL FINDINGS

- The convoy and map tasks successfully elicit reinforcement learning and cognitive flexibility.
- The transition from exploration and exploitation can be captured.
- Synchronization of disparate data streams (i.e., eye tracking, EEG, and decision data) is possible.
- The data collected have shown promise in revealing patterns for EEG and eye tracking.
- The high level of between-subject variability in decision performance speaks to the need for the proposed decision models and to its potential use to detect suboptimal decision making, due to TBI and other neurological problems.
- The combination of relatively simple measures (latency variance and regret) can indicate, in near-immediate time, the need for a training intervention for trainees whose cognitive state is misaligned with their actual decision performance.

FUTURE WORK

For the third year, we will continue to explore various statistical methods of characterizing the transition from exploration to exploitation. Results of these efforts will be submitted to peer-reviewed journals and conferences. We will also analyze the eye-tracking and EEG data from Study 2.

We will collaborate with WRIISC in determining if the convoy and map tasks can provide unobtrusive indicators of TBI status and cognitive functioning. Finally, we will transition our methodology and findings to a project funded by the Navy that aims to more effectively train recruiters. For year three, we expect to complete papers from Studies 1 and 2, and to conduct and report the results from the follow-on study designed in year two. Anticipated paper topics include:

- Correlation between neurophysiological measures and decision performance.
- Modeling human decision making on the convoy and map tasks (method of maintaining estimate, level of exploration, and level of discounting).
- Comparing performance of algorithms on convoy and map tasks.
- Assessing decision-making performance with EEG to guide training interventions.
- Comparing how decisions and underlying cognitive strategies differ when tactical leaders work with a live wingman versus an automated wingman.

APPENDIX A: CONVOY TASK PENALTY SCRIPT

Selection	Route 1	Route 2	Route 3	Route 4
1	–350	0	–50	–250
2	–250	–1,250	–50	0
3	0	0	0	0
4	–200	0	–50	0
5	0	0	0	0
6	–300	0	–50	0
7	0	0	0	0
8	–150	0	–50	0
9	0	0	0	0
10	0	0	0	0
11	–350	0	–50	–250
12	–250	–1,250	–50	0
13	0	0	0	0
14	–200	0	–50	0
15	0	0	0	0
16	–300	0	–50	0
17	0	0	0	0
18	–150	0	–50	0
19	0	0	0	0
20	0	0	0	0
21	–350	0	–50	–250
22	–250	–1,250	–50	0
23	0	0	0	0
24	–200	0	–50	0
25	0	0	0	0
26	–300	0	–50	0
27	0	0	0	0
28	–150	0	–50	0
29	0	0	0	0
30	0	0	0	0
31	–350	0	–50	–250
32	–250	–1,250	–50	0
33	0	0	0	0
34	–200	0	–50	0
35	0	0	0	0
36	–300	0	–50	0
37	0	0	0	0
38	–150	0	–50	0
39	0	0	0	0
40	0	0	0	0
41	–350	0	–50	–250
42	–250	–1,250	–50	0
43	0	0	0	0
44	–200	0	–50	0
45	0	0	0	0

46	–300	0	–50	0
47	0	0	0	0
48	–150	0	–50	0
49	0	0	0	0
50	0	0	0	0
51	–350	0	–50	–250
52	–250	–1,250	–50	0
53	0	0	0	0
54	–200	0	–50	0
55	0	0	0	0
56	–300	0	–50	0
57	0	0	0	0
58	–150	0	–50	0
59	0	0	0	0
60	0	0	0	0
61	–350	0	–50	–250
62	–250	–1,250	–50	0
63	0	0	0	0
64	–200	0	–50	0
65	0	0	0	0
66	–300	0	–50	0
67	0	0	0	0
68	–150	0	–50	0
69	0	0	0	0
70	0	0	0	0
71	–350	0	–50	–250
72	–250	–1,250	–50	0
73	0	0	0	0
74	–200	0	–50	0
75	0	0	0	0
76	–300	0	–50	0
77	0	0	0	0
78	–150	0	–50	0
79	0	0	0	0
80	0	0	0	0
81	–350	0	–50	–250
82	–250	–1,250	–50	0
83	0	0	0	0
84	–200	0	–50	0
85	0	0	0	0
86	–300	0	–50	0
87	0	0	0	0
88	–150	0	–50	0
89	0	0	0	0
90	0	0	0	0
91	–350	0	–50	–250
92	–250	–1,250	–50	0
93	0	0	0	0
94	–200	0	–50	0

95	0	0	0	0
96	–300	0	–50	0
97	0	0	0	0
98	–150	0	–50	0
99	0	0	0	0
100	0	0	0	0
101	–350	0	–50	–250
102	–250	–1,250	–50	0
103	0	0	0	0
104	–200	0	–50	0
105	0	0	0	0
106	–300	0	–50	0
107	0	0	0	0
108	–150	0	–50	0
109	0	0	0	0
110	0	0	0	0
111	–350	0	–50	–250
112	–250	–1,250	–50	0
113	0	0	0	0
114	–200	0	–50	0
115	0	0	0	0
116	–300	0	–50	0
117	0	0	0	0
118	–150	0	–50	0
119	0	0	0	0
120	0	0	0	0
121	–350	0	–50	–250
122	–250	–1,250	–50	0
123	0	0	0	0
124	–200	0	–50	0
125	0	0	0	0
126	–300	0	–50	0
127	0	0	0	0
128	–150	0	–50	0
129	0	0	0	0
130	0	0	0	0
131	–350	0	–50	–250
132	–250	–1,250	–50	0
133	0	0	0	0
134	–200	0	–50	0
135	0	0	0	0
136	–300	0	–50	0
137	0	0	0	0
138	–150	0	–50	0
139	0	0	0	0
140	0	0	0	0
141	–350	0	–50	–250
142	–250	–1,250	–50	0
143	0	0	0	0

144	–200	0	–50	0
145	0	0	0	0
146	–300	0	–50	0
147	0	0	0	0
148	–150	0	–50	0
149	0	0	0	0
150	0 –350	0	0	0
151	–250	0	–50	–250
152	0	–1,250	–50	0
153	–200	0	0	0
154	0	0	–50	0
155	–300	0	0	0
156	0	0	–50	0
157	–150	0	0	0
158	0	0	–50	0
159	0	0	0	0
160	–350	0	0	0
161	–250	0	–50	–250
162	0	–1,250	–50	0
163	–200	0	0	0
164	0	0	–50	0
165	–300	0	0	0
166	0	0	–50	0
167	–150	0	0	0
168	0	0	–50	0
169	0	0	0	0
170	–350	0	0	0
171	–250	0	–50	–250
172	0	–1,250	–50	0
173	–200	0	0	0
174	0	0	–50	0
175	–300	0	0	0
176	0	0	–50	0
177	–150	0	0	0
178	0	0	–50	0
179	0	0	0	0
180	–350	0	0	0
181	–250	0	–50	–250
182	0	–1250	–50	0
183	–200	0	0	0
184	0	0	–50	0
185	–300	0	0	0
186	0	0	–50	0
187	–150	0	0	0
188	0	0	–50	0
189	0	0	0	0
190	–350	0	0	0
191	–250	0	–50	–250
192	0	–1,250	–50	0

193	–200	0	0	0
194	0	0	–50	0
195	–300	0	0	0
196	0	0	–50	0
197	–150	0	0	0
198	0	0	–50	0
199	0	0	0	0
200	–350	0	0	0

Table 7. Script of scheduled Friendly Damage returned by route and times that route has been selected.

APPENDIX B: DEMOGRAPHIC SURVEY

Demographic Survey

Subject # Date
1. Age: _____

2. Gender: Male _____Female _____

3. What is your preferred hand for writing? Right _____Left _____

4. Do you serve or have you served in any armed forces? Yes No

5. If yes, which branch? _____ Rank: _____Years: _____

6. How many total months have you been deployed?

7. When was your most recent deployment?

8. Where was your most recent deployment?

9. During your most recent deployment, what were your main responsibilities?

****
<u>To be completed by the experimenter:</u>
Visual acuity:
Left eye_____
Right eye_____
Overall_____

APPENDIX C: POSTTASK SURVEY FORM

Subject#: Date:

Convoy Task

1. During the convoy task, how did you determine which road to select?

2. Did you use a particular strategy? If so, what was it?

3. Please rate the routes from safest (1) to most dangerous (4):

Top left road	Top right road
Bottom left road	Bottom right road

Map matching task:

1. On which map features did you sort?

2. How quickly did you realize that the sorting rule had changed? Check the response that best characterizes your overall experience.

___Immediately/After 1-2 trials
___After a few trials (3-4 trials)
___After several trials (5+ trials)
___Did not realize sorting rule had changed

Please continue to questions on back of sheet.

EEG:

1. How comfortable was it wearing the EEG cap?

2. Do you think it affected your performance on any of the tests? If so, how?

Additional comments:

Are there any additional comments for the study team?

Thank you for your participation!

APPENDIX D: STUDENT THESIS

A Comparison of Tactical Leader decision making with Automated or Live Counterparts in a Virtual Environment (Virtual Battlefield Simulation 2)

This thesis completed by Major Scott Patton examined whether tactical leaders who vary in tactical decision-making experience make different decisions when they have an automated wingman versus a live wingman (Patton, 2014). Below is the abstract. For full details, see (Patton, 2014).

THESIS ABSTRACT

The use of "responsible" autonomous systems may not be far away. Prior to developing or using responsible autonomous systems, it may be important to know if tactical leaders would make different types of decisions with automated systems than they would make with a human live crew. This work attempts to determine if decisions, time to make decisions, and confidence in decisions differ when tactical leaders rely on an autonomous wingman or a live wingman. Virtual Battlespace Simulation 2 was used to provide the virtual environment in which 30 military personnel completed a simulated mission that entailed five decision points. Subjects were randomly assigned to have an autonomous or live wingman. Decision patterns were compared to a standard based on Army Doctrine for mechanized infantry Bradley sections and subject matter experts. Results indicated no significant group difference in decisions made, time to make decisions, and confidence in decisions. However, significant group differences emerged in the aspects of the wingman that subjects trusted most and least. Although most subjects indicated that they would not trust autonomous wingmen in real combat, results suggest that subjects would revert to doctrinal decisions when faced with an unambiguous situation with an unmanned system with which they had some experience.

FUTURE DATA ANALYSES

Subjects' visual scan and brain activity was measured while they completed the tactical decision making scenario. Because of key aspects of the decision task, the combination of real-time, neurophysiological and behavioral decision data will increase our understanding of optimal wargaming decision making. The task is dynamic; captures real-world, tactical decisions; and subjects are provided with a mix of relevant and irrelevant visual information. Additionally, results will provide insight into how tactical leaders handle new technology (such as an automated wingman). For example, do they attend to the same pieces of information prior to making a decision? With these characteristics, we will be able to test the model of nonoptimal decision making depicted in Figure 11.

REFERENCES

Bechara, A., Damasio, H., & Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7–15.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa gambling task and the somatic marker hypothesis: Some questions and answers. *Trends in Cognitive Sciences*, *9*(4), 159–162.

Fricker, R. (2010). *Introduction to statistical methods for biosurveillance.* Cambridge: Cambridge University Press.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Wiegl-type card-sorting problem. *Journal of Experimental Psycholog.*, *38*(4), 404–411.

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. D. (2001). *Comparison of expert and novice scan behaviors during VFR flight*. Paper presented at the 11th International Symposium on Aviation Psychology, Columbus, OH.

Lopez, T. (2011). Odierno outlines priorites as Army chief. *Army News Service*. Retrieved from http://www.defense.gov/News/NewsArticle.aspx?ID=65292

Marshall, S. (2007). Identifying cognitive state from eye metrics. *Aviat. Space Environ. Med.*, *78*(5), B165–B175.

Nesbitt, P., Kennedy, Q., Alt, J., Fricker, R., Wihitaker, L., Yang, J. H., . . . Patton, S. (2014). Understanding optimal decision making in wargaming. Monterey, CA: Naval Postgraduate School.

Shan, I., Chen, Y., Lee, Y., & Su, T. (2008). Adult normative data on the Wisconsin card sorting test in Taiwan. *Journal of Chinese Medical Association*, *71*(10), 517–522.

Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., & Wagenmakers, E. J. (2013). Performance of healthy participants on the Iowa Gambling Task. *Psychological Assessment*, *25*(1), 180.

Sullivan, J., Yang, J. H., Day, M., & Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan patterns. *Avation*, *Space*, *and Environmental Medicine*, *82*(9), 871–878.

Vartanian, O., & Mandel, D. R. (2011). *Neuroscience of decision making*. New York City: Psychology Press.

Wechsler, D. (2008). *Weschler adult intelligence scale* (4th ed.). San Antonio, TX: Psychological Corporation.